Cosmological perturbation theory and conserved quantities in the large-scale limit

Winfried Zimdahl

Fakultät für Physik, Universität Konstanz, PF 5560 M678 D-78434 Konstanz, Germany

The linear cosmological perturbation theory of an almost homogeneous and isotropic perfect fluid universe is reconsidered and formally simplified by introducing new covariant and gauge-invariant variables with physical interpretations on hypersurfaces of constant expansion, constant curvature or constant energy density. The existence of conserved perturbation quantities on scales larger than the Hubble scale is discussed. The quantity which is conserved on large scales in a flat background universe may be expressed in terms of the fractional, spatial gradient of the energy density on constant expansion hypersurfaces or, alternatively, with the help of expansion or curvature perturbation variables on hypersurfaces of constant energy density. For nonvanishing background curvature the perturbation dynamics is most suitably described in terms of energy density perturbations on hypersurfaces of constant curvature.

I. INTRODUCTION

Understanding the dynamics of small inhomogeneities in the early universe is essential for any theory of cosmological structure formation. The conventional, linear cosmological perturbation theory relies on a splitting of all metric and matter quantities into a homogeneous and isotropic zeroth order and small, first-order perturbations about this background. This procedure implies a splitting of the spacetime itself. The nonuniqueness of this splitting is the source of the so-called gauge problem [1–4]. A perturbation analysis may either be performed in a specific gauge, the traditionally favoured gauge used to be the synchronous gauge (see, e.g., [5]), or in a gauge-invariant manner. While the choice of a specific gauge corresponds to a specific identification of the points of the fictitious background spacetime with those of the real spacetime, a gauge-invariant description is independent of this identification. Therefore, a gauge-invariant approach is conceptionally more attractive although the physical meaning of the gauge-invariant variables in terms of which the theory is formulated becomes obvious only for specific observers.

According to their extension with respect to the Hubble scale, cosmological perturbations are divided into large-scale perturbations, i.e., perturbations with wavelengths larger than the Hubble length, and small-scale perturbations for which the opposite is true.

In inflationary cosmology the presently observed structures are traced back to quantum fluctuations during an early de Sitter phase. These originally small-scale perturbations are stretched out tremendously in the inflationary period, thereby crossing the Hubble length (which is constant during the de Sitter stage) and becoming large-scale perturbations. Lateron, after the inflation has finished, i.e., when the universe is adequately described by the standard Friedmann-Lemaître-Robertson-Walker (FLRW) model, these perturbations cross the Hubble length again, this time inwards, and again become small-scale perturbations.

Consequently, it is of interest to follow the perturbation dynamics during the time interval between both Hubble scale crossings. Bardeen [1], Bardeen, Steinhardt and Turner [6] were the first to recognize the existence of a conserved quantity on large scales which simplified the corresponding perturbation analysis. The existence of a conservation quantity allows one to establish a link between physical variables in remote cosmological periods. This well-known fact has been widely used in the literature (see, e.g., [7–10]).

Recently, the role of conserved quantities was discussed in terms of the gauge-invariant and covariant Ellis-Bruni variables [3] by Dunsby and Bruni [11]. The key quantities of the Ellis-Bruni analysis are the covariantly defined spatial gradients of the energy density, the pressure, the expansion and the 3-curvature scalar. From their definition all these variables have a general physical meaning for a comoving (with the fluid 4-velocity) observer without any reference to perturbation theory. For small deviations from homogeneity and isotropy they may be related to usual perturbation variables which are gauge-invariant by construction. The relations among the Ellis-Bruni variables in linear order and the corresponding results for the cosmological modes coincide with those of the comoving gauge in conventional perturbation theory [12]. For different gauges, the Ellis-Bruni quantities, if regarded as conventional, linear perturbation quantities (see below) have no obvious physical meaning, in general. In other words, as to their physical interpretation, comoving observers are preferred.

The present paper introduces a set of covariant and gauge-invariant quantities which represent suitable combinations of Ellis-Bruni type variables. Like the original Ellis-Bruni quantities, the new variables are general tensorial expressions which may be used to characterize deviations from homogeneity and isotropy without explicitly introducing a fictitious background universe, i.e., they are no perturbation variables in the conventional sense. While the Ellis-Bruni quantities are adapted to comoving hypersurfaces, the new variables are associated to constant expansion hypersurfaces, i.e., hypersurfaces of uniform Hubble parameter, hypersurfaces of constant 3-curvature, and constant density hypersurfaces.

We shall establish the connections of these new quantities both to the Ellis-Bruni variables and among themselves. We will reconsider the perturbation theory of a perfect fluid universe and we will show that the corresponding analysis formally simplifies if written in terms of the new variables. This simplification is especialy useful with respect to the characterization of a conserved quantity in the limit of large perturbation scales.

For a flat background we find expressions for the latter in terms of the fractional spatial energy density gradient (or the spatial gradient of the 3-curvature) on constant expansion hypersurfaces, or, alternatively, in terms of the expansion (or 3-curvature) gradients on constant density hypersurfaces. In the case of nonvanishing background curvature none of these quantities will be conserved, in general. Only for dust or under the additional condition that the curvature terms in the background field equations are small, the above quantities remain approximately constant. The most suitable quantity to characterize this case turns out to be the fractional, spatial gradient of the energy density on hypersurfaces of constant 3-curvature.

The paper is organized as follows: In section II we recall the gauge-problem in conventional perturbation theory and point out the relation between the gauge-invariant perturbation variables used previously by the author [14,15] and the Ellis-Bruni variables [3]. In section III we consider the cosmological dynamics in terms of the Ellis-Bruni variables as modified by Jackson [16]. In section IV we define covariant and gauge-invariant variables associated with constant expansion hypersurfaces, i.e. hypersurfaces of constant Hubble parameter, hypersurfaces of constant curvature and constant density hypersurfaces, and we establish relations between them. In section V different expressions for large-scale conserved quantities for vanishing background curvature are found and the problems to generalize these results to curved backgrounds are discussed. Section VI presents the conclusions of the paper.

II. THE GAUGE PROBLEM

The investigation of cosmological perturbations within general relativity was pioneered by Lifshitz [17]. The traditional approach (see, e.g. [5]) starts with Einstein's field equations

$$G_{mn} \equiv R_{mn} - \frac{1}{2}g_{mn}R = \kappa T_{mn} \tag{1}$$

$(m, n.... = 0, 1, 2, 3)$ and splits both sides of the latter into a zeroth-order background and first-order perturbations about this background:

$$G_{mn} = G_{mn}^{(0)} + \hat{G}_{mn} , \quad T_{mn} = T_{mn}^{(0)} + \hat{T}_{mn} . \tag{2}$$

Provided, the zeroth-order problem $G_{mn}^{(0)} = \kappa T_{mn}^{(0)}$ is solved, one has to study the set of perturbation equations

$$\hat{G}_{mn} = \kappa \hat{T}_{mn} . \tag{3}$$

Restricting ourselves to a perfect fluid with the energy-momentum tensor

$$T_{mn} = \rho u_m u_n + p h_{mn} , \tag{4}$$

where ρ is the energy density, p is the pressure, u^m is the 4-velocity and $h^{mn} = g^{mn} + u^m u^n$, the equations (3) represent a system of differential equations for the perturbation quantities

$$\hat{\rho} \equiv \rho - \rho^{(0)} , \quad \hat{p} \equiv p - p^{(0)} , \quad \hat{u}_m \equiv u_m - u_m^{(0)} , \quad \hat{g}_{mn} \equiv g_{mn} - g_{mn}^{(0)} . \tag{5}$$

None of these quantities is gauge-invariant, i.e., invariant under infinitesimal coordinate transformations

$$x^{n\prime} = x^n - \zeta^n (x) . \tag{6}$$

Assuming a homogeneous and isotropic, comoving (i.e., $u^{m(0)} = (1, 0, 0, 0),$) zeroth order, the transformation properties of these quantities are (see, e.g., [5], chapter 10.9)

$$\hat{\rho}' = \hat{\rho} + \dot{\rho}^{(0)}\zeta^0 \ , \quad \hat{p}' = \hat{p} + \dot{p}^{(0)}\zeta^0 \ , \quad \hat{u}'_m = \hat{u}_m - \zeta^0_{,m} \ , \tag{7}$$

where we have used $\rho^{(0)}_{,0} = \dot{\rho}^{(0)} \equiv \rho^{(0)}_{,n}u^{n(0)}$ etc. in the comoving zeroth order, and

$$\hat{g}'_{mn} = \hat{g}_{mn} + \zeta^a_{,m}g^{(0)}_{an} + \zeta^a_{,n}g^{(0)}_{ma} + \zeta^a g^{(0)}_{mn,a} \ . \tag{8}$$

One may either chose now a specific gauge, e.g., the traditionally preferred synchronous gauge, or try to find a gauge-invariant desciption ([1]). In the latter case one may look for suitable combinations of the quantities (5) that are gauge-invariant, i.e., invariant under the infinitesimal coordinate transformations (6). From the transformation properties (7) it is obvious that, e.g., the quantity

$$\hat{\rho}^{(c)}_{,\mu} \equiv \hat{\rho}_{,\mu} + \dot{\rho}^{(0)}\hat{u}_\mu \tag{9}$$

($\mu... = 1, 2, 3$) is invariant under the transformations (6), i.e. $\hat{\rho}^{(c)\prime}_{,\mu} = \hat{\rho}^{(c)}_{,\mu}$. This quantity represents the energy density perturbations on comoving (superscript 'c') hypersurfaces $\hat{u}_\mu = 0$ (cf [14,15]). It corresponds to the spatial derivative of Bardeen's quantity ε_m [1,13].

One easily recognizes that this construction principle of gauge-invariant quantities is applicable for any scalar ([14,15]). E.g., pressure perturbations on comoving hypersurfaces are characterized by

$$\hat{p}^{(c)}_{,\mu} \equiv \hat{p}_{,\mu} + \dot{p}^{(0)}\hat{u}_\mu \ . \tag{10}$$

Gauge-invariant descriptions of scalars such as the fluid expansion $\Theta \equiv u^a_{;a}$ and the spatial 3-curvature $\mathcal{R}$, to be defined below, are

$$\hat{\Theta}^{(c)}_{,\mu} = \hat{\Theta}_{,\mu} + \dot{\Theta}^{(0)}\hat{u}_\mu \ , \quad \hat{\mathcal{R}}^{(c)}_{,\mu} = \hat{\mathcal{R}}_{,\mu} + \dot{\mathcal{R}}^{(0)}\hat{u}_\mu \ , \tag{11}$$

respectively. All these quantities are gauge-invariant by construction. In order to establish a link between these variables and those introduced by Ellis and Bruni [3], we have to investigate the key quantities of the latter authors, the spatially projected gradients $h^c_a\rho_{,c}$, $h^c_ap_{,c}$, $h^c_a\Theta_{,c}$, $h^c_a\mathcal{R}_{,c}$, etc., in first order. Using the first-order expressions

$$\hat{h}^\alpha_\beta = 0 \ , \quad \hat{h}^0_\alpha = u_\alpha \ , \tag{12}$$

one finds

$$(h^c_0\rho_{,c})\hat{} = 0 \ , \quad (h^c_\mu\rho_{,c})\hat{} = \hat{\rho}_{,\mu} + \dot{\rho}^{(0)}\hat{u}_\mu \equiv \hat{\rho}^{(c)}_{,\mu} \ , \tag{13}$$

and corresponding relations for the spatial gradients of p, Θ, and $\mathcal{R}$ in first order. While the quantities $h^c_a\rho_{,c}$ etc. vanish in zeroth order since the background was assumed to be homogeneous, they coincide in first order with the gauge-invariant quantities $\hat{\rho}^{(c)}_{,\mu}$ etc.

The second relation (13) demonstrates explicitly how the Ellis-Bruni variables are related to corresponding gauge-invariant perturbation quantities with physical interpretations on comoving hypersurfaces. Our considerations may be regarded as a kind of motivation for the usefulness of the Ellis-Bruni quantities from the point of view of conventional perturbation theory. They also establish the connection beween the latter quantities and the quantities of the type (9) and (10), introduced in [14,15]. Using the Ellis-Bruni quantities which, by their definition, vanish in a homogeneous universe, one may avoid an explicit decomposition into a fictitious background and perturbations about this (not unique) zeroth order. Since the existence of a gauge-problem is just a consequence of the nonuniqueness of the background, one circumvents the gauge-problem from the beginning by using the Ellis-Bruni variables which are covariant by definition and no perturbation quantities in the usual sense.

The spatially projected gradients by Ellis and Bruni have a physical meaning on comoving hypersurfaces. As to the physical understanding of the gauge-invariant quantities $h^c_a\rho_{,c}$ etc., if regarded as conventional first-order perturbation variables in the sense described above, the comoving gauge is naturally preferred. In different gauges, these quantities have no obvious physical meaning, in general. From these statements the impression may arise that there is something special about the comoving gauge, even within a gauge-invariant formalism. This is not the case, however, as we are going to clarify in this paper. It is well-known that there exist obviously reasonable gauge-invariant quantities in the literature having a simple physical meaning in gauges different from the comoving gauge. For reasons of convenience some authors e.g., Bardeen, Steinhardt and Turner [6], occasionally prefer the 'uniform Hubble constant gauge' or others, e.g., Hwang [22], the constant curvature gauge. One may therefore ask whether there exist covariant and gauge-invariant quantities that, from the point of view of their physical interpretation, prefer hypersurfaces of constant Hubble parameter or of constant curvature in a similar sense in which the Ellis-Bruni quantities prefer comoving hypersurfaces. As we will see below, there are indeed problems for which a choice of covariant variables different from the Ellis-Bruni choice may simplify the dynamical description. In order to introduce these variables we first recall the basic elements of the cosmological dynamics in terms of the Ellis-Bruni quantities, following here the elegant presentation by Jackson [16].

III. COSMOLOGICAL DYNAMICS IN TERMS OF THE ELLIS-BRUNI-JACKSON VARIABLES

Instead of applying the field equations (1), we shall investigate the cosmological dynamics within the 'fluid-flow' approach used in [18,19,7,20,21]. The equations of motion $T^{ik}_{;k} = 0$, imply

$$\dot{\rho} = -\Theta\left(\rho + p\right) \;, \tag{14}$$

with

$$\Theta \equiv u^i_{\;;i} \;, \tag{15}$$

and

$$\left(\rho + p\right)\dot{u}^m = -p_{,k}h^{mk} \;, \tag{16}$$

where $\dot{u}^m \equiv u^m_{;n}u^n$. Additionally, we shall use the Raychaudhuri equation for Θ,

$$\dot{\Theta} + \frac{1}{3}\Theta^2 + 2\left(\sigma^2 - \omega^2\right) - \dot{u}^a_{;a} - \Lambda + \frac{\kappa}{2}\left(\rho + 3p\right) = 0 \;. \tag{17}$$

Λ is the cosmological constant. The magnitudes of shear and vorticity are defined by

$$\sigma^2 \equiv \frac{1}{2}\sigma_{ab}\sigma^{ab} \;, \qquad \omega^2 \equiv \frac{1}{2}\omega_{ab}\omega^{ab} \;, \tag{18}$$

with

$$\sigma_{ab} = h^c_a h^d_b u_{(c;d)} - \frac{1}{3}\Theta h_{ab} \;, \qquad \omega_{ab} = h^c_a h^d_b u_{[c;d]} \;. \tag{19}$$

The 3-curvature scalar of the projected metric,

$$\mathcal{R} = 2\left(-\frac{1}{3}\Theta^2 + \sigma^2 - \omega^2 + \kappa\rho + \Lambda\right) \;, \tag{20}$$

reduces to the 3-curvature of the surfaces orhogonal to u^a in the case $\omega = 0$. The homogeneous and isotropic FLRW universes are characterized by $\sigma = \omega = \dot{u}^a = 0$. Taking the spatial gradient of (14) yields

$$h^c_a\dot{\rho}_{,c} = -\left(\rho + p\right)h^c_a\Theta_{,c} - \Theta h^c_a\left(\rho + p\right)_{,c} \;. \tag{21}$$

Introducing a length scale S by

$$\frac{1}{3}\Theta \equiv \frac{\dot{S}}{S} \tag{22}$$

and rewriting the l.h.s. of (21) after multiplying by S as

$$Sh^c_m\dot{\rho}_{,c} = h^a_m\left(Sh^c_a\rho_{,c}\right)^{\cdot} - \Theta Sh^c_m p_{,c} + \left(\omega^c_m + \sigma^c_m\right)Sh^n_c\rho_{,n} \;, \tag{23}$$

the equations (21) and (23) may be combined into

$$h^a_n\dot{D}_a + \frac{\dot{p}}{\rho + p}D_n + \left(\omega^c_n + \sigma^c_n\right)D_c + t_n = 0 \;, \tag{24}$$

where we have introduced the fractional quantity [16]

$$D_a \equiv \frac{Sh^c_a\rho_{,c}}{\rho + p} \;, \tag{25}$$

and

$$t_a \equiv Sh^c_a\Theta_{,c} \;. \tag{26}$$

4

Equation (24) is Jackson's eq.(29) [16]. Similarly, taking the spatially projected gradient of the Raychaudhuri equation (17), one obtains

$$h_n^a \dot{t}_a = -\dot{\Theta} P_n - (\omega_n^c + \sigma_n^c) t_c - \frac{2}{3}\Theta t_n - Sh_n^c \left(2\sigma^2 - 2\omega^2\right)_{,c}$$

$$+ Sh_n^c \left(\dot{u}_{;a}^a\right)_{,c} - \frac{\kappa}{2}(\rho + p)\left[D_n + 3P_n\right] , \tag{27}$$

with the fractional quantity

$$P_a \equiv \frac{Sh_a^c p_{,c}}{\rho + p} , \tag{28}$$

characterizing the pressure perturbations. Equation (27) is Jackson's eq.(30) [16]. The set of equations (24) and (27) is still completely general. Even for an equation of state $p = p(\rho)$ which allows one to express P_n in terms of D_n, the equations (24) and (27) are, of course, not a closed system for D_a and t_a since these quantities are coupled to ω and σ and their spatial gradients.

From now on we shall assume the spatial gradients as well as σ and ω to be small, i.e., we assume the universe to be almost homogeneous and isotropic. Consequently, up to first order in the inhomogeneities, the factors in front of the quantities D_a, P_a and t_a in (24) and (27) refer to the homogeneous and isotropic case with $\omega = \sigma = 0$. The linearized set of equations becomes

$$h_n^a \dot{D}_a + \frac{\dot{p}}{\rho + p} D_n + t_n = 0 , \tag{29}$$

and

$$h_n^a \dot{t}_a = -\frac{1}{2}\mathcal{R} P_n - \frac{2}{3}\Theta t_n - \frac{\kappa}{2}(\rho + p) D_n + Sh_n^c \left(\dot{u}_{;a}^a\right)_{,c} , \tag{30}$$

where we have used the zeroth-order relations ($\Lambda = 0$)

$$\kappa\rho = \frac{1}{3}\Theta^2 + \frac{1}{2}\mathcal{R} , \tag{31}$$

and

$$\dot{\Theta} + \frac{3}{2}\kappa(\rho + p) = \frac{1}{2}\mathcal{R} , \tag{32}$$

for homogeneous and isotropic universes. The 3-curvature in the latter case is known to be

$$\mathcal{R} = \frac{6k}{a^2} , \tag{33}$$

with the scale factor a of the Robertson-Walker metric. The last term in (30) is generally given by ([16])

$$Sh_n^c \left(\dot{u}_{;a}^a\right)_{,c} = -Sh_n^c \left[h^{ab}\left(h_a^m \frac{p_{,m}}{\rho + p}\right)_{;b}\right]_{,c} + Sh_n^c \left[h^{ab}\frac{p_{,a}}{\rho + p}\frac{p_{,b}}{\rho + p}\right]_{,c} . \tag{34}$$

In linear order in the inhomogeneities this reduces to

$$Sh_n^c \left(\dot{u}_{;a}^a\right)_{,c} = -\frac{\nabla^2}{a^2} P_n . \tag{35}$$

With an equation of state $p = p(\rho)$ one may write

$$\dot{p} = -c_s^2 \Theta(\rho + p) , \qquad P_n = c_s^2 D_n , \tag{36}$$

where $c_s^2 \equiv (\partial p/\partial\rho)_{ad}$ is the square of the sound velocity. Because of (13) etc., in linear order, the system (29), (30) may be written as

$$\dot{D}_\mu - c_s^2 \Theta D_\mu + t_\mu = 0 , \tag{37}$$

5

and

$$\dot{t}_\mu = -\frac{2}{3}\Theta t_\mu - \left[\frac{\kappa}{2}\left(\rho + p\right) + 3\frac{k}{a^2}c_s^2 + c_s^2\frac{\nabla^2}{a^2}\right]D_\mu \ . \tag{38}$$

Eliminating t_μ from the system (37), (38), one obtains a closed equation for D_μ,

$$\ddot{D}_\mu + \left(\frac{2}{3} - c_s^2\right)\Theta\dot{D}_\mu$$
$$- \left[\left(c_s^2\right)^{\cdot}\Theta + \frac{\kappa}{2}\left(\rho - 3p\right)c_s^2 + \frac{\kappa}{2}\left(\rho + p\right) + c_s^2\frac{\nabla^2}{a^2}\right]D_\mu = 0 \ , \tag{39}$$

which corresponds to Jackson's eqation (57). From the latter equation one easily derives the well-known growing and decaying modes in the long-wavelength limit for $k = 0$, equivalent to neglecting the spatial gradient term in the bracket in front of D_μ.

IV. NEW PERTURBATION VARIABLES

Let us consider the ratio of the spatial variation of a scalar quantity, say, the energy density, i.e., $h_a^c\rho_{,c}$ to its variation in time, $\dot{\rho} \equiv u^c\rho_{,c}$. With (13) we find, in linear order,

$$\left(\frac{h_\mu^c\rho_{,c}}{u^c\rho_{,c}}\right)^{\hat{}} = \frac{\hat{\rho}_{,\mu}}{\dot{\rho}} + \hat{u}_\mu \ . \tag{40}$$

Obviously, this ratio is a reasonable quantity to characterize small deviations from homogeneity. Similar ratios may be formed for the other scalar quantities of interest:

$$\left(\frac{h_\mu^c p_{,c}}{u^c p_{,c}}\right)^{\hat{}} = \frac{\hat{p}_{,\mu}}{\dot{p}} + \hat{u}_\mu \ , \tag{41}$$

and

$$\left(\frac{h_\mu^c\Theta_{,c}}{u^c\Theta_{,c}}\right)^{\hat{}} = \frac{\hat{\Theta}_{,\mu}}{\dot{\Theta}} + \hat{u}_\mu \ . \tag{42}$$

For $k \neq 0$ we may also introduce

$$\left(\frac{h_\mu^c\mathcal{R}_{,c}}{u^c\mathcal{R}_{,c}}\right)^{\hat{}} = \frac{\hat{\mathcal{R}}_{,\mu}}{\dot{\mathcal{R}}} + \hat{u}_\mu \qquad (k \neq 0) \ . \tag{43}$$

For $k = \mathcal{R} = 0$ the 3-curvature perturbations are gauge-invariant, i.e.,

$$\left(h_\mu^c\mathcal{R}_{,c}\right)^{\hat{}} = \hat{\mathcal{R}}_{,\mu} \qquad (k = 0) \ . \tag{44}$$

By construction, all the quantities (40) - (43) represent the first-order ratio of the spatial variation of the quantities ρ, p, Θ and $\mathcal{R}$, respectively, to their change in time, on comoving hypersurfaces $\hat{u}_\mu = 0$. The 4-velocity perturbation $\hat{u}_\mu$ enters each of the relations (40) - (43) in exactly the same manner, namely simply additively. Forming differences between any two of the quantities (40) - (43), the velocity perturbations just cancel. This suggests combining (40) - (43) in order to obtain new, covariant and gauge-invariant variables. E.g., we may define

$$\frac{h_m^c\rho_{,c}^{(ce)}}{u^c\rho_{,c}} \equiv \frac{h_m^c\rho_{,c}}{u^c\rho_{,c}} - \frac{h_m^c\Theta_{,c}}{u^c\Theta_{,c}} \ , \qquad (k = 0, \pm 1) \ , \tag{45}$$

or

$$\frac{h_m^c\rho_{,c}^{(cc)}}{u^c\rho_{,c}} \equiv \frac{h_m^c\rho_{,c}}{u^c\rho_{,c}} - \frac{h_m^c\mathcal{R}_{,c}}{u^c\mathcal{R}_{,c}} \qquad (k = \pm 1) \ . \tag{46}$$

In linear order, with (40), (42) and (43), the spatial components are

$$\left(\frac{h_\mu^c \rho_{,c}^{(ce)}}{u^c \rho_{,c}}\right)^{\widehat{\ }} = \frac{\hat{\rho}_{,\mu}}{\dot{\rho}} - \frac{\hat{\Theta}_{,\mu}}{\dot{\Theta}} \qquad (k = 0, \pm 1) \ , \tag{47}$$

and

$$\left(\frac{h_\mu^c \rho_{,c}^{(cc)}}{u^c \rho_{,c}}\right)^{\widehat{\ }} = \frac{\hat{\rho}_{,\mu}}{\dot{\rho}} - \frac{\hat{\mathcal{R}}_{,\mu}}{\dot{\mathcal{R}}} \qquad (k = \pm 1) \ . \tag{48}$$

Generally, all variables of the type (45) and (46) are exact covariant quantities without any reference to perturbation theory. As perturbation variables they are gauge-invariant by construction. From (47) it seems obvious to conclude that $h_m^c \rho_{,c}^{(ce)} / (u^c \rho_{,c})$ in first order represents the ratio of spatial to time variation of the energy density on hypersurfaces $\hat{\Theta}_{,\mu} = 0$, i.e., on hypersurfaces of constant expansion or constant Hubble parameter, in the same sense as the Ellis-Bruni quantity $h_m^c \rho_{,c} / (u^c \rho_{,c})$ has a corresponding, well-known meaning in the comoving gauge $\hat{u}_\mu = 0$ (cf.(13)). While we shall indeed use this interpretation of (47) in the following, it requires additional clarification. The point is that the quantity ρ, introduced by (4), is the energy density for a comoving (with 4-velocity u^a) observer. A different observer, moving, e.g., with a 4-velocity n^a, normal to hypersurfaces $\hat{\Theta}_{,\mu} = 0$, would interpret a different quantity, namely $\mu = T_{ab} n^a n^b$ as energy density. (Different from the observer moving with u^a he would also measure an energy flux.) Generally, ρ and μ are related by (see [23,13])

$$\mu = \rho \cosh^2 \beta + p \sinh^2 \beta \ ,$$

where $\beta(t)$ is the hyperbolic angle of tilt given by $\cosh \beta = -u^a n_a$. Corresponding relations hold for the other quantities like pressure and expansion. It follows that the perturbation quantity $\hat{\rho}$ for a comoving (with u^a) observer will generally not coincide with the perturbation $\hat{\mu}$, i.e., $\hat{\rho}$ will not coincide with the energy density perturbation on hypersurfaces of constant expansion. Now, in the present case we are considering first-order perturbations, while u^a and n^a coincide in zeroth order. For small angles of tilt, however, i.e., for $\beta \ll 1$, the differences between μ and ρ are of second order in β. Consequently, within linear perturbation theory we may identify the quantities $\hat{\mu}$ and $\hat{\rho}$. Similar statements hold for the other perturbation variables. Therefore, the above interpretation of regarding the quantity (47) as the ratio of spatial to time change of the energy density on hypersurfaces of constant expansion, $\hat{\Theta}_{,\mu} = 0$, is justified up to first order. Since the latter gauge is defined by the vanishing of the gradient of $\hat{\Theta}$, the quantity $\hat{\Theta}$ itself is only determined up to a constant.

Analogously, for $k = \pm 1$, the quantity $h_m^c \rho_{,c}^{(cc)} / (u^c \rho_{,c})$ in first order represents the ratio of spatial to time change of the energy density on hypersurfaces of constant curvature (superscript 'cc'), defined by $\hat{\mathcal{R}}_{,\mu} = 0$. The same construction principle may be applied to introduce a variable describing the ratio of spatial to time change of the 3-curvature,

$$\frac{h_m^c \mathcal{R}_{,c}^{(ce)}}{u^c \mathcal{R}_{,c}} \equiv \frac{h_m^c \mathcal{R}_{,c}}{u^c \mathcal{R}_{,c}} - \frac{h_m^c \Theta_{,c}}{u^c \Theta_{,c}} \qquad (k = \pm 1) \ . \tag{49}$$

For $k = 0$ one has $h_\mu^c \mathcal{R}_{,c}^{(ce)} = h_\mu^c \mathcal{R}_{,c}$.

The ratio of spatial to time change of the expansion with respect to the corresponding ratio for the 3-curvature is

$$\frac{h_m^c \Theta_{,c}^{(cc)}}{u^c \Theta_{,c}} = \frac{h_m^c \Theta_{,c}}{u^c \Theta_{,c}} - \frac{h_m^c \mathcal{R}_{,c}}{u^c \mathcal{R}_{,c}} \qquad (k = \pm 1) \ . \tag{50}$$

Again, the 0-component vanishes and the first-order spatial components may be written analogously to (47) and (48), namely

$$\left(\frac{h_\mu^c \mathcal{R}_{,c}^{(ce)}}{u^c \mathcal{R}_{,c}}\right)^{\widehat{\ }} = \frac{\hat{\mathcal{R}}_{,\mu}}{\dot{\mathcal{R}}} - \frac{\hat{\Theta}_{,\mu}}{\dot{\Theta}} \qquad (k = \pm 1) \ , \tag{51}$$

and

$$\left(\frac{h_\mu^c \Theta_{,c}^{(cc)}}{u^c \Theta_{,c}}\right)^{\widehat{\ }} = \frac{\hat{\Theta}_{,\mu}}{\dot{\Theta}} - \frac{\hat{\mathcal{R}}_{,\mu}}{\dot{\mathcal{R}}} \qquad (k = \pm 1) \ . \tag{52}$$

The quantity (51) represents the first-order ratio of spatial to time change of the 3-curvature on constant expansion hypersurfaces, while (52) is the corresponding ratio for the expansion on constant curvature hypersurfaces.

Provided, an equation of state $p = p(\rho)$ is given, we are left with three quantities being ratios of spatial to time changes, namely those for ρ, Θ and $\mathcal{R}$. Forming differences between them yields variables that in linear order may be interpreted as perturbation variables, e.g., on constant expansion hypersurfaces like (47) and (51), or, like (48) and (52), on hypersurfaces of constant curvature. In a similar way it is possible to introduce variables which in first order have physical interpretations on constant density hypersurfaces. The corresponding definitions are

$$\frac{h_m^c \mathcal{R}_{,c}^{(cd)}}{u^c \mathcal{R}_{,c}} \equiv \frac{h_m^c \mathcal{R}_{,c}}{u^c \mathcal{R}_{,c}} - \frac{h_m^c \rho_{,c}}{u^c \rho_{,c}} \qquad (k = \pm 1) \ , \tag{53}$$

and

$$\frac{h_m^c \Theta_{,c}^{(cd)}}{u^c \Theta_{,c}} = \frac{h_m^c \Theta_{,c}}{u^c \Theta_{,c}} - \frac{h_m^c \rho_{,c}}{u^c \rho_{,c}} \qquad (k = 0, \pm 1) \ , \tag{54}$$

where the superscript 'cd' stands for 'constant density'. In linear order one has

$$\left(\frac{h_\mu^c \mathcal{R}_{,c}^{(cd)}}{u^c \mathcal{R}_{,c}} \right)^{\hat{}} = \frac{\hat{\mathcal{R}}_{,\mu}}{\dot{\mathcal{R}}} - \frac{\hat{\rho}_{,\mu}}{\dot{\rho}} \qquad (k = \pm 1) \ , \tag{55}$$

and

$$\left(\frac{h_\mu^c \Theta_{,c}^{(cd)}}{u^c \Theta_{,c}} \right)^{\hat{}} = \frac{\hat{\Theta}_{,\mu}}{\dot{\Theta}} - \frac{\hat{\rho}_{,\mu}}{\dot{\rho}} \qquad (k = 0, \pm 1) \ . \tag{56}$$

The quantity (55) is (48) with the opposite sign. A corresponding relation holds between (56) and (47).

The relations of all of these quantities to the Ellis-Bruni variables are obvious. It will turn out that the perturbation dynamics looks simpler in any of the new variables than it does in terms of the Ellis-Bruni variables. The reason is that the new variables are adapted to hypersurfaces on which one of the first-order spatial gradients of Θ, $\mathcal{R}$ or ρ vanishes. None of these gradients vanishes, however, on comoving hypersurfaces.

A. Constant expansion hypersurfaces

Differentiating the Gauß-Codazzi equation (20) and projecting orthogonal to u^a yields, in linear order in the inhomogeneities,

$$r_a = -\frac{4}{3} \Theta t_a + 2\kappa (\rho + p) D_a \ , \tag{57}$$

where we have introduced the abbreviation

$$r_a \equiv S h_a^c \mathcal{R}_{,c} \ . \tag{58}$$

According to (45) and (47) we introduce the fractional quantity

$$D_a^{(ce)} \equiv D_a - \frac{\dot{\rho}}{\rho + p} \frac{t_a}{\dot{\Theta}} \qquad (k = 0, \pm 1) \quad , \tag{59}$$

which in linear order represents the fractional, spatial gradient of the energy density on hypersurfaces of constant expansion. Analogously, we define the fractional spatial gradient of the curvature on constant expansion hypersurfaces,

$$r_a^{(ce)} \equiv r_a - \dot{\mathcal{R}} \frac{t_a}{\dot{\Theta}} \qquad (k = \pm 1) \ , \tag{60}$$

with (cf (33))

$$\dot{\mathcal{R}} = -\frac{2}{3} \Theta \, \mathcal{R} \ . \tag{61}$$

Using the definitions (59) and (60) to replace D_a and t_a, respectively, in (57), we find a relation between the gradients of the 3-curvature and the fractional energy density on uniform Hubble parameter hypersurfaces:

$$r_a^{(ce)} = 2\kappa \left(\rho + p\right) D_a^{(ce)} , \qquad (k = 0, \pm 1) . \qquad (62)$$

This is relation (57) written in the variables that are adapted to constant expansion hypersurfaces. The latter relation is also valid for $k = 0$, where $r_a^{(ce)} \overset{(k=0)}{=} r_a$.

B. Constant curvature hypersurfaces

According to (50) and (52) the fractional energy density gradient on hypersurfaces of constant curvature is characterized by

$$D_a^{(cc)} \equiv D_a - \frac{\dot{\rho}}{\rho + p} \frac{r_a}{\dot{\mathcal{R}}} \qquad (k = \pm 1) . \qquad (63)$$

Analogously, for the gradient of the expansion on these hypersurfaces one has

$$t_a^{(cc)} \equiv t_a - \dot{\Theta} \frac{r_a}{\dot{\mathcal{R}}} \qquad (k = \pm 1) . \qquad (64)$$

Inserting D_a and t_a from (63) and (64) into (57), and using the zeroth-order relations (32) and (61), we get

$$\frac{4}{3} \Theta t_a^{(cc)} = 2\kappa \left(\rho + p\right) D_a^{(cc)} \qquad (k = \pm 1) , \qquad (65)$$

which is again (57), now in terms of variables adapted to constant curvature hypersurfaces.

Equation (57) is a linear relation between the three Ellis-Bruni type perturbation quantities D_a, t_a and r_a defined with respect to comoving hypersurfaces. The introduction of the variables (59) and (60) or, (63) and (64), respectively, reduces (57) to relations (either (62) or (65)) between only two variables, adapted either to hypersurfaces of constant expansion or to those of constant curvature. Effectively, the number of variables has been reduced in both cases. One may choose, e.g., one of the fractional energy density perturbations, $D_a^{(cc)}$ or $D_a^{(ce)}$ as independent variable.

At any stage it is possible to change between the sets of variables associated with either the comoving, or the uniform Hubble constant, or the constant curvature hypersurfaces. Subtracting (63) and (59) yields

$$D_a^{(cc)} - D_a^{(ce)} = -\frac{\dot{\rho}}{\rho + p} \frac{r_a^{(ce)}}{\dot{\mathcal{R}}} \qquad (k = \pm 1) . \qquad (66)$$

Using here the relation (62) between $r_a^{(ce)}$ and $D_a^{(ce)}$ as well as the zeroth-order relations (32) and (61), we find the following connection between the fractional energy density gradients on constant curvature and constant expansion hypersurfaces:

$$\frac{1}{2} \mathcal{R} D_a^{(cc)} = \dot{\Theta} D_a^{(ce)} \qquad (k = \pm 1) . \qquad (67)$$

C. Constant density hypersurfaces

We define, according to (54) and (56) with (26), the perturbations of the expansion on constant density hypersurfaces,

$$t_a^{(cd)} \equiv t_a - \left(\rho + p\right) \dot{\Theta} \frac{D_a}{\dot{\rho}} \qquad (k = 0, \pm 1) , \qquad (68)$$

as well as, according to (53) and (55) with (58), curvature perturbations on these hypersurfaces,

$$r_a^{(cd)} \equiv r_a - \left(\rho + p\right) \dot{\mathcal{R}} \frac{D_a}{\dot{\rho}} \qquad (k = \pm 1) . \qquad (69)$$

9

Solving eq.(68) for t_a, eq.(69) for r_a and inserting into (57) yields

$$r_a^{(cd)} = -\frac{4}{3}\Theta t_a^{(cd)} \qquad (k = 0, \pm 1) \tag{70}$$

for the perturbed Gauß-Codazzi equation in terms of variables that are adapted to constant density hypersurfaces. The relation (70) is also valid for $k = 0$, where $r_a^{(cd)} \stackrel{(k=0)}{=} r_a$. Introducing into (68) the quantity $D_a^{(ce)}$ by (59), one obtains a relation between the quantities $t_a^{(cd)}$ and $D_a^{(ce)}$,

$$\Theta t_a^{(cd)} = \dot{\Theta} D_a^{(ce)} \qquad (k = 0, \pm 1) . \tag{71}$$

Analogously, from (69) and (63) one gets

$$r_a^{(cd)} = -\frac{2}{3}\mathcal{R} D_a^{(cc)} \qquad (k = \pm 1) . \tag{72}$$

It follows, that $t_a^{(cd)}$ and $D_a^{(cc)}$ are connected through

$$\Theta t_a^{(cd)} = \frac{1}{2}\mathcal{R} D_a^{(cc)} \qquad (k = \pm 1) . \tag{73}$$

V. CONSERVED QUANTITIES

We are now going to demonstrate that the system (37), (38) becomes especially simple if rewritten in terms of the newly introduced variables. Starting with the definition (63) of the fractional energy density gradient on hypersurfaces of constant curvature and using (14) and (61), we have

$$D_a^{(cc)} = D_a - \frac{3}{2}\frac{r_a}{\mathcal{R}} \qquad (k = \pm 1) . \tag{74}$$

Introducing here (57) for r_a and applying the zeroth-order relation (32) yields

$$D_a^{(cc)} = \frac{2}{\mathcal{R}}\left[\dot{\Theta} D_a + \Theta t_a\right] \qquad (k = \pm 1) . \tag{75}$$

Differentiating the latter relation and using the equations of the system (37) and (38) on its r.h.s. provides us with

$$\dot{D}_\mu^{(cc)} = -\frac{\dot{p}}{\rho + p}D_\mu - \Theta P_\mu - \frac{2}{\mathcal{R}}\Theta\frac{\nabla^2}{a^2}P_\mu \qquad (k = \pm 1) . \tag{76}$$

According to (36) the first and second terms on the r.h.s. of (76) cancel and we arrive at

$$\dot{D}_\mu^{(cc)} = -\frac{2}{\mathcal{R}}\Theta\frac{\nabla^2}{a^2}c_s^2 D_\mu \qquad (k = \pm 1) . \tag{77}$$

This simple relation comprises the entire linear perturbation dynamics for nonflat background universes. In the case $k = 0$, where the 'cc'-quantities are not defined, one has to modify the above reasoning to arrive at a corresponding relation. From the definition (59) one finds with (14),

$$\dot{\Theta} D_a^{(ce)} = \dot{\Theta} D_a + \Theta t_a \qquad (k = 0, \pm 1) . \tag{78}$$

The r.h.s. of this equation coincides with the expression in the bracket on the r.h.s. of (75). After similar steps like those between eqs.(75) and (77) one gets

$$\left[a^2\dot{\Theta} D_\mu^{(ce)}\right]^{\cdot} = -a^2\Theta c_s^2\frac{\nabla^2}{a^2}D_\mu \qquad (k = 0 \pm 1) . \tag{79}$$

Alternatively, taking into account (62), (71) and (70), the quantity $a^2\dot{\Theta} D_\mu^{(ce)}$ may be replaced to yield

$$\left[a^2 \dot{\Theta} \frac{r_\mu^{(ce)}}{2\kappa \, (\rho + p)}\right]^{\cdot} = \left[a^2 \Theta t_\mu^{(cd)}\right]^{\cdot} = -\frac{3}{4} \left[a^2 r_\mu^{(cd)}\right]^{\cdot} = -a^2 \Theta c_s^2 \frac{\nabla^2}{a^2} D_\mu \; , \tag{80}$$

for $k = 0 \pm 1$. For $k \neq 0$ equation (79) coincides with (77). For $k = 0$ the quantities $a^2 r_a^{(ce)}$ and $a^2 r_a^{(cd)}$ reduce to the variable C_a used by Dunsby and Bruni [11].

Equation (79) (or an equivalent formulation according to (80) or (77)) represents the most condensed form of the linear perturbation dynamics. It is the main result of this paper. A comparison with (39) shows that all terms in that equation, with the exception of the spatial gradient term, have been included into a first time derivative. The price to pay for this formal simplicity is that in general each of the equations (77), (79) and (80) couples variables that have physical interpretations in different gauges.

It is obvious that the introduction of the new variables is especially useful in cases where the r.h.s. of (79) and (80) (or (77)) may be neglected.

For $D_\mu = D_{(n)} \nabla_\mu Q_{(n)}$ (and corresponding relations for $D_\mu^{(ce)}$ etc.) where the $Q_{(n)}$ satisfy the Helmholtz equation $\nabla^2 Q_{(n)} = -n^2 Q_{(n)}$ one has ([24,25,2,11]) $n^2 = \nu^2$ for $k = 0$ and $n^2 = \nu^2 + 1$ for $k = -1$, where ν is continous and related to the physical wavelength by $\lambda = 2\pi a / \nu$. For $k = +1$ the eigenvalue spectrum is discrete, namely $n^2 = m\,(m+2)$ with $m = 1, 2, 3.....$

On large perturbation scales, i.e., for $\nu \ll 1$ in a flat background universe the r.h.s. of (79) and (80) may be neglected and we find

$$a^2 \dot{\Theta} D_{(\nu)}^{(ce)} = a^2 \dot{\Theta} \frac{r_{(\nu)}^{(ce)}}{2\kappa \, (\rho + p)} = a^2 \Theta t_{(\nu)}^{(cd)} = -\frac{3}{4} a^2 r_{(\nu)}^{(cd)} \approx const \qquad (\nu \ll 1) \tag{81}$$

for $k = 0$.

Using equation (59) which defines $D_a^{(ce)}$ in terms of D_μ and t_μ, and eliminating t_μ by (37), yields the following relation between $D_a^{(ce)}$ and D_a:

$$\dot{\Theta} D_{(\nu)}^{(ce)} = -\Theta \dot{D}_{(\nu)} + \left[\left(c_s^2 - \frac{1}{3}\right)\Theta^2 - \frac{\kappa}{2}\,(\rho + 3p)\right] D_{(\nu)} \qquad (\nu \ll 1) \; . \tag{82}$$

Of course, the latter relation and (79) are equivalent to (39). On large scales and for $k = 0$ equation (82) together with (81) is a first order equation to determine $D_{(\nu)}$.

In a nonflat universe (for simplicity we restrict ourselves to the case $k = -1$ which is observationally favoured) the r.h.s. of (77), (79) and (80) do not generally vanish even for $\nu \ll 1$. The only exception is $c_s^2 \ll 1$, i.e., a dust universe (cf [11]). Using the above eigenvalue structure of the Laplacian in (39) and comparing the last term in the bracket in front of D_μ in (39) with the previous ones by taking into account the zeroth-order equation (31), one finds that the spatial gradient term may be neglected for $\nu \ll 1$, provided the additional condition

$$H^2 a^2 \gg 1 \tag{83}$$

with $H \equiv \Theta/3$ is satisfied. The latter relation, however, coincides with the condition under which the curvature term in the background equations (31) and (32) may be neglected. Consequently, the quantities $D_\mu^{(cc)}$, $a^2 \dot{\Theta} D_\mu^{(ce)}$ and equivalent expressions under the dot derivative on the l.h.s. of (80) are only conserved on large scales if the background curvature terms are negligible. Corresponding properties for a differently defined quantity were found by Dunsby and Bruni [11](see eq(16) in [11]. For special cases these authors, however, also constructed conservation quantities without the restriction (83) and on arbitrary scales.)

While the quantities $D_\mu^{(ce)}$, $r_\mu^{(ce)}$, $t_\mu^{(cd)}$ and $r_\mu^{(cd)}$ which appear on the l.h.s. of (79) and (80) are well defined for any value of the background curvature, the quantity $D_\mu^{(cc)}$, the fractional energy density perturbation on hypersurfaces of constant curvature, defined only for $k \neq 0$, appears to be the most suitable quantity to characterize perturbations on a nonflat background, however small the curvature terms may be. The general perturbation dynamics reduces to (77) in this case and $D_\mu^{(cc)}$ is conserved on large scales if $a^2 H^2 \gg 1$.

VI. CONCLUSIONS

We have simplified the linear, perfect fluid cosmological perturbation theory by the introduction of new, covariant and gauge-invariant perturbation variables. These variables are physically interpreted on hypersurfaces of constant

expansion, constant 3-curvature, or constant energy density. Their relations to the Ellis-Bruni variables which have physical interpretations on comoving hypersurfaces are established. The conserved quantity in the large-scale limit for $k = 0$ may be expressed in terms of the gradient of the fractional energy density on constant expansion hypersurfaces or, alternatively, in terms of the expansion gradient on constant density hypersurfaces. The most suitable quantity for the description of the perturbation dynamics in a universe with nonvanishing background curvature is the covariantly defined fractional, spatial gradient of the energy density on hypersurfaces of constant 3-curvature. It is this quantity which is approximately conserved on large scales provided the background curvature terms are small.

One may hope that use of the variables introduced in the present paper also simplifies perturbation calculations under more general circumstances.

[1] Bardeen J M 1980 *Phys.Rev.* **D22** 1882
[2] Kodama H and Sasaki M 1984 *Prog.Theor.Phys.Suppl.* **78** 1
[3] Ellis G F R and Bruni M 1989 *Phys.Rev.* **D40** 1804
[4] Stewart J M 1990 *Class.Quantum Grav.* **7** 1169
[5] Weinberg S 1972 *Gravitation and Cosmology* (New York: Wiley)
[6] Bardeen J M, Steinhardt P and Turner M 1983 *Phys.Rev.* **D28** 679
[7] Lyth D 1985 *Phys.Rev.* **D31** 1792
[8] Hwang J 1991 *Astrophys.J.* **380** 307
[9] Mukhanov V F, Feldman H A and Brandenberger R H 1992 *Phys.Rep.* **215** 203
[10] Liddle A R and Lyth H l 1993 *Phys.Rep.* **231** 1
[11] Dunsby P K S and Bruni M 1994 *Int.J.Mod.Phys.* **D3** 443
[12] Ellis G F R, Hwang J and Bruni M 1989 *Phys.Rev.* **D40** 1819
[13] Bruni M, Dunsby PKS and Ellis G F R 1992 *Astrophys.J.* **395** 34
[14] Zimdahl W 1991 *Class.Quantum Grav.* **8** 677
[15] Zimdahl W 1992 *Astrophys.J.* **393** 471
[16] Jackson J C 1993 *Mon.Not.R.Astron.Soc.* **264** 729
[17] Lifshitz E M 1946 *J.Phys.*(USSR) **10** 116
[18] Hawking S W 1966 *Astrophys.J.* **145** 544
[19] Olson D W 1976 *Phys.Rev.* **14** 327
[20] Lyth D H and Mukherjee M 1988 *Phys.Rev.* **38** 485
[21] Lyth D H and Stewart E D 1990 *Astrophys.J.* **361** 343
[22] Hwang J 1993 *Phys.Rev.* **D48** 3544
[23] King A R and Ellis G F R 1973 *Comm.Math.Phys.* **31** 209
[24] Lifshitz E M and Khalatnikov I M 1963 *Adv.Phys.* **12** 185
[25] Harrison E R 1967 *Rev.Mod.Phys.* **39** 862